DONALD DAVIDSON

QUOTATION

Quotation is a device used to refer to typographical or phonetic shapes by exhibiting samples, that is, inscriptions or utterances that have those shapes. This characterization is broad and vague: broad enough to include not only written quotation marks, and spoken phrases like "and I quote," but also the finger-dance quotes often used by philosophers condemned to read aloud what they have written; and vague enough to leave open the question whether the words that began this sentence ("This characterization") show a form of quotation.

In quotation not only does language turn on itself, but it does so word by word and expression by expression, and this reflexive twist is inseparable from the convenience and universal applicability of the device. Here we already have enough to draw the interest of the philosopher of language; but one discerns as well connections with further areas of concern such as sentences about propositional attitudes, explicit performatives, and picture theories of reference. If the problems raised by quotation appear trivial by comparison, we may welcome finding an easy entrance to the labyrinth.

When I was initiated into the mysteries of logic and semantics, quotation was usually introduced as a somewhat shady device, and the introduction was accompanied by a stern sermon on the sin of confusing the use and mention of expressions. The connection between quotation on the one hand and the use–mention distinction on the other is obvious, for an expression that would be used if one of its tokens appeared in a normal context is mentioned if one of its tokens appears in quotation marks (or some similar contrivance for quotation). The invitation to sin is perhaps accounted for by the ease with which quotation marks may be overlooked or omitted. But the strictures on quotation often sound a darker note. Thus Tarski, in "The Concept of Truth in Formalized Languages," examines the possibilities for an articulate theory of quotation marks, and decides that one is led at once to absurdities, ambiguities and contradiction.[1] Quine wrote in *Mathematical*

Logic, "Scrupulous use of quotation marks is the main practical measure against confusing objects with their names . . ." but then he adds that quotation

> . . . has a certain anomalous feature which calls for special caution: from the standpoint of logical analysis each whole quotation must be regarded as a single word or sign, whose parts count for no more than serifs or syllables. A quotation is not a *description*, but a *hieroglyph*; it designates its object not by describing it in terms of other objects, but by picturing it. The meaning of the whole does not depend upon the meanings of the constituent words.[2]

And Church, while praising Frege for his careful use of quotation to avoid equivocation, himself eschews quotation as "misleading", "awkward in practice . . . and open to some unfortunate abuses and misunderstandings".[3] There is more than a hint, then, that there is something obscure or confused about quotation. But this can't be right. There is nothing wrong with the device itself. It is our theories about how it works that are inadequate or confused.

It is often said that in quotation, the quoted expressions are mentioned and not used. The first part of this claim is relatively clear. It is the second part, which says quoted expressions aren't used, that seems suspect. Why isn't incorporation into quotation one use of an expression? A plausible response would be that of course there is *some* sense in which the quoted material is used, but its use in quotation is unrelated to its *meaning* in the language; so the quoted material is not used as a piece of language.

This response may not quite still our doubts. For one thing, there are the troublesome cases where it is convenient both to use and to mention the same expression by speaking or inscribing a single token of the expression. I once resolved to adopt a consistent way of using quotation in my professional writing. My plan was to use single quotation marks when I wanted to refer to the expression a token of which was within, but double quotation marks when I wanted to use the expression in its usual meaning while at the same time indicating that the word was odd or special ("scare quotes"). I blush to admit that I struggled with this absurd and unworkable formula for a couple of years before it dawned on me that the second category contained the seeds of its own destruction. Consider, for example, a passage earlier in this paper where I say, nearly enough:

> Quine says that quotation ". . . has a certain anomalous feature."

Are the quoted words used or mentioned? Obviously mentioned, since the words are Quine's own, and I want to mark the fact. But equally obvious is the fact that the words are used; if they were not, what follows the word "quotation" would be a singular term, and this it cannot be if I have produced a grammatical sentence. Nor is it easy to rephrase my words so as to resolve the difficulty. For example, it is not enough to write, "Quine used the words 'has a certain anomalous feature' of quotation," for this leaves out what he meant by those words.

Here is another mixed case of use and mention that is not altogether easy to sort out:

> Dhaulagiri is adjacent to Anapurna, the mountain whose conquest Maurice Herzog described in his book of the same name.

The last phrase "the same name" cannot mean the same name as the mountain, for the mountain has many names. Rather it means the same name of the mountain as the one used earlier in the sentence. I would call this a genuine case of quotation, for the sentence refers to an expression by exhibiting a token of that expression; but it is a case that manages without quotation marks.

Or consider this case:

> The rules of Clouting and Dragoff apply, in that order.[4]

Temporarily setting aside these last examples as pathological and perhaps curable, there is a way, now standard, of giving support to the idea that in quotation the quoted material is not used. This is the interpretation of quotation proposed by Tarski as the only one he can defend. According to it a quotation, consisting of an expression flanked by quotation marks, is like a single word and is to be regarded as logically simple. The letters and spaces in the quoted material are viewed as accidents in the spelling of a longer word and hence as meaningless in isolation. A quotation mark name is thus, Tarski says, like the proper name of a man.[5] I shall call this the *proper name theory* of quotation. Church attributes the same idea, or at least a method with the same consequences, to Frege. Church writes:

Frege introduced the device of systematically indicating autonomy by quotation marks, and in his later publications (though not in the *Begriffsschrift*) words and symbols used autonomously are enclosed in single quotation marks in all cases. This has the effect that

a word enclosed in single quotation marks is to be treated as a different word from that
without the quotation marks — as if the quotation marks were two additional letters in
the spelling of the word — and equivocacy is thus removed by providing two different
words to correspond to the different meanings.[6]

Unless I am mistaken, this passage exhibits a common confusion. For what
expression is it, according to the view Church attributes to Frege, that refers
to the word a token of which appears inside the quotation marks? Is it that
word itself (given the context), or the quotation as a whole? Church says
both, though they cannot be identical. The word itself, since an expression in
quotation marks has a meaning distinct from its usual meaning; it is "treated
as a different word" which is used "autonomously." The quotation as a
whole, since the quotation marks are part of the spelling.

Quine has repeatedly and colorfully promoted the idea of the quotation as
unstructured singular term. Not only is there his denial, already cited, that
quotations are descriptions, but the claim that the letters inside the quotation
marks in a quotation occur there ". . . merely as a fragment of a longer name
which contains, beside this fragment, the two quotation marks."[7]

The merit in this approach to quotation is the emphasis it puts on the fact
that the reference of a quotation cannot be construed as owed, at least in any
normal way, to the reference (or meaning) of the expressions displayed
within the quotation marks. But it seems to me that as an account of how
quotation works in natural language, the approach is radically deficient. If
quotations are structureless singular terms, then there is no more significance
to the *category* of quotation mark names than to the category of names that
begin and end with the letter 'a' ("Atlanta," "Alabama," "Alta," "Athena,"
etc.). On this view, there is no relation, beyond an accident of spelling,
between an expression and the quotation mark name of that expression. If
we accept this theory, nothing would be lost if for each quotation mark name
we were to substitute some unrelated name, for that is the character of
proper names. And so no echo remains, as far as this theory of quotation
goes, of the informal rules governing quotation that seem so clear: if you
want to form a quotation mark name of an expression, flank that expression
with quotation marks; and, a quotation mark name refers to "its interior"
(as Quine puts it). Nothing left, either, of the intuitively attractive notion
that a quotation somehow pictures what it is about.

These objections are in themselves enough to throw doubt on Tarski's

claim that this interpretation of quotation is ". . . the most natural one and completely in accordance with the customary way of using quotation marks . . . ."[8] But there is a further and, I think, decisive objection, which is that on this theory we cannot give a satisfactory account of the conditions under which an arbitrary sentence containing a quotation is true. In an adequate theory, every sentence is construed as owing its truth or falsity to how it is built from a finite stock of parts by repeated application of a finite number of modes of combination. There are, of course, an infinite number of quotation mark names, since every expression has its own quotation mark name, and there are an infinite number of expressions. But on the theory of quotation we are considering, quotation mark names have no significant structure. It follows that a theory of truth could not be made to cover generally sentences containing quotations. We must reject the proper name interpretation of quotation if we want a satisfactory theory for a language containing quotations.

I turn now to a quite different theory of quotation, which may be called the *picture theory* of quotation. According to this view, it is not the entire quotation, that is, expression named plus quotation marks, that refers to the expression, but rather the expression itself. The role of the quotation marks is to indicate how we are to take the expression within: the quotation marks constitute a linguistic environment within which expressions do something special. This was perhaps the view of Reichenbach, who said that quotation marks ". . . transform a sign into a name of that sign."[9] Quine also suggests this idea when he writes that a quotation ". . . designates its object . . . by picturing it,"[10] for of course it is only the interior of a quotation that could be said to be like the expression referred to (the quotation marks are not in the picture — they are the frame). And Church also, in the passage just discussed, toys with the notion that on Frege's theory ". . . a word enclosed in single quotation marks is to be treated as a different word" in that it is used "autonomously," that is, to name itself.

It should be allowed at once that the three authors just mentioned, in the passages alluded to, vacillate between the proper name theory of quotation and the picture theory. Yet the theories are clearly distinct; so bearing in mind the deficiencies of the proper name theory, we ought to consider the picture theory on its own. At first sight it promises two advantages: it attributes *some* structure to quotations, since it treats them as composed of

quotation marks (which set the scene for interpreting their contents) and the quoted material. And it hints, in its appeal to the relation of picturing, at a theory that will draw on our intuitive understanding of how quotation works.

These seeming advantages fade when examined. The difficulty is this. What is wanted is an explanation of how quotation enables us to refer to expressions by picturing them. But on the present theory, quotation marks create a context in which expressions refer to themselves. How then does picturing feature in the theory? If an expression inside quotation marks refers to itself, the fact that it also pictures itself is simply a diverting irrelevancy.

Would it help to say that quotation marks create a context in which we are to view the contents as a picture of what is referred to? Not at all; this is merely a tendentious way of saying the expression refers to itself. In brief, once the content of the quotation is assigned a standard linguistic role, the fact that it happens to resemble something has no more significance for semantics than onomatopoeia or the fact that the word "polysyllabic" is polysyllabic.

Another important point might escape us here. The picturing relation as between an object and itself is hardly interesting and the theory, as we are interpreting it, tries vainly to make something of this drab idea. But the more interesting picturing we sense in quotation is not between expression and expression. In quotation, what allows us to refer to a certain expression, which we may take to be an abstract shape, is the fact that we have before us on the page or in the air something that *has* that shape – a token, written or spoken. The picture theory suggests no way to bring an inscription or utterance into the picture. This could be done only by describing, naming, or pointing out the relevant *token*, and no machinery for the purpose has been introduced.

The picture theory of quotation is reminiscent of Frege's theory of opaque (what he called oblique) contexts such as those created by "necessarily", "Jones believes that . . .", "Galileo said that . . . ", and so on. There are conspicuous differences between these contexts as analyzed by Frege, and quotation as treated by the picture theory: in quotation words may change their part of speech (since every expression becomes a name or description) while in the other contexts this never happens; and in quotation, but not in other opaque contexts, nonsense makes sense. But there is the striking similarity that in both cases some linguistic device is supposed to create a context

within which words play new referential roles. This concept of a context that alters reference has never been properly explained, and Frege himself was leery of it: it certainly does not lend itself to direct treatment in a theory of truth. The trouble with the picture theory, as with Frege's treatment of opaque contexts generally, is that the references attributed to words or expressions in their special contexts are not functions of their references in ordinary contexts, and so the special context-creating expressions (like quotation marks or the words "said that") cannot be viewed as functional expressions.[11]

A central defect of the proper name theory of quotation was that while viewing quotations as well-formed expressions of the language, it failed to provide an articulate theory showing how each of the infinitude of such expressions owed its reference to its structure. The experiment just concluded showed that it is possible to treat quotations as having semantically significant structure. Let us press on in this direction.

Geach has long insisted that quotations are really *descriptions*, and hence have structure, and he complains of the proper name theory as I have[12] (though he does not connect his complaints with the need for a theory of truth). His theory, as I understand it, is this. A single word in quotation marks names itself; this is a new item of vocabulary, and is not semantically complex (I am not sure whether Geach says this last). So far, the theory is like the proper name theory. But a longer expression when quoted is a structured description. Thus " "Alice swooned" " abbreviates " "Alice" "Swooned" " which reads "the expression got by writing " "Alice" " followed by " "swooned" ". This theory has the advantages of the preceding Fregean theory, and is far simpler and more natural. (It may be called the *spelling theory* of quotation.)

Both Tarski and Quine imply, by things they say, that they see the possibility of a similar theory. Thus Tarski remarks that if we accept the name theory, then quotation mark names can be eliminated and replaced everywhere by structural-descriptive names,[13] while Quine contends that we can dispel the opacity of quotation, when we please, by resorting to spelling.[14] The device both have in mind is like Geach's except that Geach takes the smallest units to be words, while Tarski and Quine take them to be individual letters and symbols. The result, in the abbreviations of ordinary quotation, is the same. In primitive notation, which reveals all structure to

the eye, Geach has an easier time writing (for only each word needs quotation marks) but a harder time learning or describing the language (he has a much larger primitive vocabulary — twice normal size).

There is no difficulty about extending a truth definition to the devices of spelling suggested by Quine, Tarski and Geach; yet these devices can be thought of as merely abbreviated by ordinary quotation. This claim of mere abbreviation may be backed by describing a mechanical method for going back and forth between the two styles of notation. Thus given the quotation mark name " "Alice swooned" ", the machine starts at the left by reproducing the first quotation marks, then the letter "A", then another set of quotation marks, then a sign for concatenation, then another set of quotation marks, and so on until it reaches a set of quotation marks in the original. It reproduces these and stops. The result will be:

$$\overparen{\text{"A"}\ \text{"l"}\ \text{"i"}\ \text{"c"}\ \text{"e"}\text{space}\text{"s"}\ \text{"w"}\ \text{"o"}\ \text{"o"}\ \text{"n"}\ \text{"e"}\ \text{"d"}}$$

Since the two notations are mechanically interchangeable, there is no reason not to consider a semantics for one a semantics for the other: so this *could* be regarded as a theory of how quotation works in English (modifications would work for other languages). But would it be a correct theory of ordinary quotation? There are several reasons for saying it would not.

Notice first that the appearance of quotation marks in the expanded notation is adventitious. The theory works by identifying a finite set of units (words or letters) from which every expression in the language to be described is composed. Then unstructured proper names of these units are introduced, along with a notation for concatenation. Such a theory works as well, and is less misleading, if quotation marks are dropped entirely and new names of the building blocks are introduced. To illustrate (following Geach's method), suppose the word "Alice" is named by the word "alc" and the word "swooned" by the word "sw"; then "Alice swooned" would be described by:

$$\overparen{\text{alc}\ \text{sw}}$$

or, using Quine's method:

$$\overparen{\text{Ay}\ \text{ell}\ \text{eye}\ \text{see}\ \text{ee}}\ \text{space}\ \overparen{\text{es}\ \text{double-you}\ \text{oh}\ \text{oh}\ \text{en}\ \text{ee}\ \text{dee}}$$

This tiny exercise is meant to emphasize the fact that nothing of the idea of quotation *marks* is captured by this theory — nothing of the idea that one can

form the name of an arbitrary expression by enclosing it in quotation marks. On the spelling theory, no articulate item in the vocabulary corresponds to quotation marks, and so the theory cannot reflect a rule for their use. The machine simply knows by heart the name of each smallest expression. Clearly, one essential element in the idea that quotations picture what they are about has been lost.

A striking way to see what is and what is not relevant to structure is to try applying existential generalization and substitution of identity. A standard way of demonstrating that quotation as normally used does not wear its structure on its surface is to observe that from:

> "Alice swooned" is a sentence

we cannot infer:

> $(\exists x)$ ("x swooned" is a sentence).

Nor, supposing "alc" names "Alice", can we infer:

> "alc swooned" is a sentence

nor:

> alc $\frown$ "swooned" is a sentence.

But (using Geach's version of the spelling theory) we *can* go from:

> "Alice swooned" is a sentence

to:

> "Alice" $\frown$ "swooned" is a sentence

and thence to:

> alc $\frown$ "swooned" is a sentence

and then to:

> $(\exists x)$ (alc $\frown$ x is a sentence)

or:

> $(\exists x)$ (x $\frown$ "swooned" is a sentence).

In Quine's version of the theory, we could go from:

> "Alice" is a word

to:

"A" "l" "i" "c" "e" is a word

to:

$(\exists x)(\exists y)(x$ "l" "i" y "e" is a word$)$.

These derivations show clearly that quotation marks play no vital role in the spelling theory; and also that this theory is not a theory of how quotation works in natural language.

One essential element of picturing has been lost, but not perhaps quite all, for the spelling theory does appear to depend on having the description of complex expressions reproduce the *order* of the expressions described. In the description provided by the theory, names of particular expressions need not resemble what they name, but in the description as a whole, names of expressions that are concatenated are themselves concatenated.

Even this residue to the picturing idea is superficial, however. The descriptions the spelling theory provides are themselves, from the point of view of a fully articulate language, mere abbreviations of something more complicated in which the order of expressions may well be changed. I think we should conclude that the spelling theory of quotation has no connection with the view that we understand quotations as picturing expressions.

There are further important uses of quotation in a natural language that cannot be explained by the spelling theory and could not be accommodated by a language constructed in the way it suggests. The spelling theory cannot, at least in any obvious way, deal with those mixed cases of use and mention we discussed earlier, nor indeed with any case that seems to depend on a demonstrative reference to an utterance or inscription. An important use for quotation in natural language is to introduce new notation by displaying it between quotation marks; this is impossible on the spelling theory provided the new notation is not composed of elements that have names. On the spelling theory we also could not use quotation to teach a foreign language based on a new alphabet or notation, for example Khmer or Chinese. Since these are functions easily performed by ordinary quotation (whether or not with quotation marks), we cannot accept the spelling theory as giving an adequate account of quotation in natural language.

We have discovered a short list of conditions to be satisfied by a competent theory of quotation. The first is that like a theory for any aspect of a

language it should merge with a general theory of truth for the sentences of the language. The other conditions are specific to quotation. One is that the theory provide an articulate semantic role for the *devices* of quotation (quotation marks, or verbal equivalents). When we learn to understand quotation we learn a rule with endless applications: if you want to refer to an expression, you may do it by putting quotation marks around a token of the expression you want to mention. A satisfactory theory must somehow embody or explain this piece of lore. And finally, a satisfactory theory must explain the sense in which a quotation pictures what is referred to, otherwise it will be inadequate to account for important uses of quotation, for example, to introduce novel pieces of notation and new alphabets.

It is not hard to produce a satisfactory theory once the requirements are clear. The main difficulty springs, perhaps it is now obvious, from the simultaneous demands that we assign articulate structure to quotations and that they picture what they mention. For articulate linguistic structure here must be that of description, and describing seems to forestall the need to picture. The call for structure is derived from the underlying demand for a theory of meaning, here thought of as a theory of truth; all that is needed is enough structure to implement the recursive characterization of a truth predicate. Still, enough structure will be too much as long as we regard the quoted material as part of the semantically significant syntax of a sentence. The cure is therefore to give up this assumption.

It is natural to assume that words that appear between the boundaries of a sentence are legitimate parts of the sentence; and in the case of quotations, we have agreed that the words within quotation marks help us to refer to those words. Yet what I propose is that those words within quotation marks are not, from a semantical point of view, part of the sentence at all. It is in fact confusing to speak of them as words. What appears in quotation marks is an *inscription*, not a shape, and what we need it for is to help refer to its shape. On my theory, which we may call the *demonstrative theory* of quotation, the inscription inside does not refer to anything at all, nor is it part of any expression that does. Rather it is the quotation marks that do all the referring, and they help refer to a shape by pointing out something that has it. On the demonstrative theory, neither the quotation as a whole (quotes plus filling) nor the filling alone is, except by accident, a singular term. The singular term is the quotation marks, which may be read "the expression a

token of which is here." Or, to bring out the way in which picturing may now be said genuinely to be involved: "the expression with the shape here pictured."

It does not discredit this theory to say that it neglects the fact that the quoted material is syntactically part of the sentence; taken in abstraction from semantics, the question of location is trivial. In spoken sentences, temporal sequence plays the role of linear arrangement in writing. But if I say "I caught a fish this big" or "I caught this fish today," my hands, or the fish, do not become part of the language. We could easily enough remove the quoted material from the heart of the sentence. Quotation is a device for pointing to inscriptions (or utterances) and can be used, and often is, for pointing to inscriptions or utterances spatially or temporally outside the quoting sentence. So if I follow a remark of yours with "Truer words were never spoke," I refer to an expression, but I do it by way of indicating an embodiment of those words in an utterance. Quotation marks could be warped so as to remove the quoted material from a sentence in which they play no semantic role. Thus instead of:

> "Alice swooned" is a sentence

we could write:

> Alice swooned.   The expression of which this is a token is a
> sentence.

Imagine the token of "this" supplemented with fingers pointing to the token of "Alice swooned."

I take it to be obvious that the demonstrative theory assigns a structure to sentences containing quotations that can be handled in a straightforward way by a theory of truth — assuming of course that there is a way of accommodating demonstratives at all, and on this point, I have already tried to indicate why there is not any real difficulty in making room for demonstrative or indexical elements in a formal theory of truth.[15] Finally, it is obvious that the picturing feature of quotation has been exploited and explained. So the demonstrative theory also authorizes the use of quotation in introducing new bits of typography and discussing languages with new alphabets. I conclude by considering how it fares with the mixed cases of use and mention on exhibit earlier.

I said that for the demonstrative theory the quoted material was no part, semantically, of the quoting sentence. But this was stronger than necessary or desirable. The device of pointing can be used on whatever is in range of the pointer, and there is no reason why an inscription in active use can't be ostended in the process of mentioning an expression. I have already indicated an important sort of case, and there are many more. ("You pay attention to what I'm going to say." "Why did you use those words?" etc.) Any token may serve as target for the arrows of quotation, so in particular a quoting sentence may after all by chance contain a token with the shape needed for the purposes of quotation. Such tokens then do double duty, once as meaningful cogs in the machine of the sentence, once as semantically neutral objects with a useful form. Thus:

> Quine says that quotation ". . . has a certain anomalous feature."

may be rendered more explicitly:

> Quine says, using words of which these are a token, that quotation has a certain anomalous feature.

(Here the "these" is accompanied by a pointing to the token of Quine's words.) As for Anapurna:

> Dhaulighiri is adjacent to Anapurna, the mountain whose conquest Maurice Herzog described in his book with a name that is this shape (pointing to the token of "Anapurna").

Finally:

> The rules of Clouting and Dragoff apply, in the order in which these tokens appear (pointing to the tokens of "Clouting" and "Dragoff").

University of Chicago

NOTES

[1] Alfred Tarski, 'The Concept of Truth in Formalized Languages', in *Logic, Semantics, Metamathematics*, Oxford, 1956, pp. 159–162.
[2] W. V. Quine, *Mathematical Logic*, Harvard, 1940, Chapter 4.
[3] Alonzo Church, *Introduction to Mathematical Logic*, Vol. 1, Princeton, 1956, Chapter 8.

[4] The example is from John Robert Ross, 'Metalinguistic Anaphora', *Linguistic Inquiry*, 1, 1970, p. 273.

[5] *Op. cit.*, p. 159.

[6] *Op. cit.*, pp. 61–62.

[7] W. V. Quine, *From a Logical Point of View*, 2nd ed., Harvard, 1961, p. 140. Compare W. V. Quine, *Methods of Logic*, Holt, 1950, p. 38 and *Word and Object*, M.I.T., 1960, p. 143; also Benson Mates, *Elementary Logic*, Oxford, 1965, p. 24.

[8] *Op. cit.*, p. 160.

[9] Hans Reichenbach, *Elements of Symbolic Logic*, Macmillan, 1947, p. 335. Reichenbach says other things that point to a different theory.

[10] *Mathematical Logic*, Chapter 4.

[11] For further discussion of this point, see my 'On Saying That', *Synthese*, 19 (1968), pp. 130–146.

[12] P. T. Geach, *Mental Acts*, Routledge and Kegan Paul, 1957, p. 79ff. Compare 'Quotation and Quantification', in *Logic Matters*, Blackwell, 1972, pp. 205–209.

[13] *Op. cit.*, p. 160.

[14] *Word and Object*, p. 212.

[15] Donald Davidson, 'Truth and Meaning', *Synthese*, 17, 1967.